CHAPTER 5

COLOR-TEXTURE SEGMENTATION USING IAKFCM-EMD

5.1 COLOR TEXTURE BASED IMAGE SEGMENTATION

CBIR is a technique for retrieving images on the basis of features such as color, texture, shape and pattern generation. Shape description is one of the important features in CBIR. This shape is obtained based on the segmentation of the combined color and texture features. Initially, the noiseless input image is converted into Lab color space. The statistical color feature such as mean and mode are extracted from lab color space. The fuzzy texture unit is determined by the extraction of local texture information from each pixel. The combined feature extraction of color and texture are segmented using Improved Adaptive Kernelized FCM strategy (IAKFCM). Finally, refinement processes are used to eliminate the misclassified pixels produced by clustering. It is based on Earth Mover Distance (EMD). The performance results have been proved in terms of Root Mean Square Error (RMSE), Pearson correlation coefficient ($r$) and Structural SIMilarity (SSIM). These metrics are used to compare the output of the segmented results with Human labeled (Ground Truth) Segmentation. This IAKFCM-EMD technique has produced better results than the K-means clustering.
5.2 IAKFCM-EMD BASED SEGMENTATION

The proposed method segments the input image based on color-texture feature extraction. Color is one common feature used in image segmentation. It is often used in conjunction with texture information to achieve better segmentation results (Ahmed & Iftekharuddin 2011). The noiseless input image is transformed from RGB to Lab color space. Lab color space is much more intuitive than RGB. Lab color system approximates human vision and its Luminance component L and the chrominance component a, b matches closely with human perception of light. It has perceptual color fidelity. It is very convenient to measure small color difference while the RGB color space does not. It can be used to make accurate color balance corrections. The color features were extracted from L and a component based on mean and mode. The ‘b’ component doesn’t produce significant color composition when compared to that of ‘a’. The combined color and texture features are used in IAKFCM and Earth Mover distance. The block diagram of IAKFCM-EMD color image segmentation and work flow of the proposed method is shown in Figure 5.1 and Figure 5.2.

![Block diagram of IAKFCM-EMD based color-texture image segmentation](image)

Figure 5.1  Block diagram of IAKFCM-EMD based color-texture image segmentation
Figure 5.2 Work Flow of the IAKFCM-EMD based segmentation process

The input image is preprocessed using the wavelet based fuzzy filter to enhance the input image. The image is converted from RGB to Lab color space in order to extract the color statistical features by calculating the mean and mode of L and a component. The texture features are extracted using fuzzy texture spectrum. The extracted features are segmented with IAKFCM and followed by EMD refinement process to have an efficient segmentation.
5.3 PREPROCESSING

Additive noise removal is one of the important steps of preprocessing. Additive noise occurs when a value from a certain distribution is added to each image pixel, e.g., Gaussian distribution. In this preprocessing stage, wavelet-based fuzzy filter is used to remove the additive noise.

5.4 FEATURE EXTRACTION

Features (color, texture, shape, and so on) are defined as a function of one or more measurements, each of which quantifies some significant characteristics of the object (Koelstra et al 2010). The collection of the features of the content is known as a feature vector. Feature vectors describe particular characteristics of an image based on the nature of the extraction method. The dimension of the vector is determined by the number of features extracted. Based upon the literature survey, the features can be classified as general and domain. General features are application-independent features such as color, texture, and shape or boundary. According to the abstraction level, they can be further classified into three categories, i) Pixel-level features are calculated at each pixel, e.g., color. ii) Local features are calculated over the results of subdivision of the image band on image segmentation or edge detection and iii) Global features are calculated over the entire image or just regular sub-area of an image. Domain-specific features are application-dependent features such as human faces, fingerprints, and conceptual features. These features are often a synthesis of low-level features for a specific domain.

Color feature usually preserves boundaries, but it is more sensitive to local color variations. Hence, color-based segmentation provides sharp edges, but often in homogeneous regions. On the other hand, texture features provide fuzzy boundaries but more in homogeneous regions. In this research
work the color (L and ‘a’ components) and texture features are combined and segmented using an IAKFCM.

5.4.1 Color Feature Extraction

Color is a widely used important feature for image representation. This is very important as it is invariant with respect to scaling, translation and rotation of an image. Color features are easy to obtain. Color can be used to differentiate images based on their features. (Agaian et al 2007). It is one of the most used features in image retrieval. The colors are described by their color space: RGB, LAB, LUV, HSV etc. RGB is the best known color space and it’s is commonly used for visualization.

In this chapter CIE Lab color space is selected to extract color features. First, the input image is converted to the CIE Lab color space, prior to the calculation of the statistical color features of mean and mode (Xoa & Tao 2010). The mean of 3x3 neighborhood pixels is the average arithmetic intensity value, computed by adding them and dividing the result by their total number (Abdullah-Al-wadud et al 2007). The statistical mean value of an image is calculated based on overlapped 3x3 window size. A statistical term ‘mode’ that refers to the most frequently occurring number found in a set of numbers. The results for extracting the color feature are shown in Figure 5.3 and 5.4. Figure ‘a’ shows noiseless input image. Figure ‘b’ shows the lab color image.
Figure 5.3  Results of color features extraction a) Noiseless input images, b) Lab color images
5.4.2 Texture Spectrum Method

Texture is another important property of images. Textures are represented by texels. The texture spectrum is introduced and described in detail by DC. He and L. Wang. They are considered to be the founders of this method. The identification of texture spectrum in an image is achieved by the extraction of local texture information for each pixel. Texture spectrum has
been introduced in the year of 1990. The texture image can be decomposed into a set of essential small units, called Texture Units (TU). The occurrence distribution of Texture Units is called Texture Spectrum (TS). The basic concept is that a texture image can be considered as a set of essential small units termed as texture units, which characterize the local texture information for a given pixel and its neighborhood (Jiji et al 2010). In a square-raster digital image, each pixel is surrounded by eight neighboring pixels. The local texture information for a pixel can be extracted from a neighborhood of 3x3 pixels, which represents the smallest complete unit (in the sense of having eight directions surrounding the pixel). The ordering of elements and representation of texture unit are shown in Figure 5.5 and 5.6.

<table>
<thead>
<tr>
<th>$E_1$</th>
<th>$E_2$</th>
<th>$E_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_8$</td>
<td></td>
<td>$E_4$</td>
</tr>
<tr>
<td>$E_7$</td>
<td>$E_6$</td>
<td>$E_5$</td>
</tr>
</tbody>
</table>

**Figure 5.5 Fuzzy Texture Unit**

<table>
<thead>
<tr>
<th>$v_i$</th>
<th>$v_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>$v_o$</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

**Figure 5.6 Texture unit representation**
A neighborhood of 3x3 pixels is denoted by a set containing nine elements: \( V = \{V_0, V_1, V_2, ..., V_8\} \) where \( V_0 \) represents the intensity value of the central pixel and \( V_i (1 \leq i \leq 8) \) the intensity value of each neighboring pixel (Ki Tae et al, 2009). Then the corresponding Texture Unit can be represented by \( TU = \{E_1, E_2, ..., E_8\} \).

\[
E_i = \begin{cases} 
0 & \text{if } V_i < V_0 \text{ and } V_i \leq X \\
1 & \text{if } V_i < V_0 \text{ and } V_i > X \\
2 & V_i = V_0 \text{ for } i = 1, 2, ..., 8 \\
3 & \text{if } V_i > V_0 \text{ and } V_i \leq Y \\
4 & \text{if } V_i > V_0 \text{ and } V_i > Y 
\end{cases}
\]

where \( X, Y \) are user-specified values.

As each element of \( TU \) has one of the three possible values, the combination of all the eight elements results in \( 3^8 = 6561 \) possible texture units in total. The 6561 texture units are labeled by using the following Equations (5.1) and (5.2):

\[
N_{TU} = \sum_{i=1}^{8} E_i \times 5^{(\frac{E_i}{2})} \quad (5.1)
\]

\[
N_{TU} \in \{0, 1, ..., 6561\} \quad (5.2)
\]

Here NTU represents the Texture Unit Number and \( E_i \) is the \( i \)th element of texture unit set, \( TU = \{E_1, E_2, ..., E_8\} \). The size of the window depends on the nature of the texture image. The Texture Unit is given in Figure 5.7 and 5.8. For example

\[
\begin{array}{ccc}
80 & 110 & 130 \\
150 & 120 & 180 \\
90 & 120 & 200 \\
\end{array}
\]

**Figure 5.7 3x3 Neighborhood**
The sample image of the color fuzzy texture spectrum is shown in Figure 5.9.

$$V = \{80, 110, 130, 180, 200, 120, 90, 150\}$$

$$TU = \{1, 3, 4, 4, 2, 0, 4\}$$

The combined feature extraction of color and texture are implemented using IAKFCM strategy. The widely used FCM for image segmentation has some limitations. This is due to the limitation in its squared-norm distance. It measures the similarity between centers and data objects of
images which are corrupted due to heavy noise and outliers. Further clustering performance is affected by the initial center.

To overcome the above limitations IAKFCM (Cao et al 2012) algorithm is proposed for image segmentation. Here Euclidean distance in objective function is replaced by kernel function (Kannan et al 2010). Using properties of kernel functions to inner product in the original space tries to map the space into higher dimensional feature space. This proposed algorithm avoids solving large differential equations and gives much faster computational speed. The algorithm is shown in section 4.2.5. The results of IAKFCM is shown in Figure 5.10.

![Figure 5.10 Results of IAKFCM](image-url)
5.6 REFINEMENT PROCESS

In this refinement process, cross-bin metric such as Earth Mover’s Distance (EMD) is used to enhance the segmentation process. The main aim of this work is to segment the given dynamic scene into foreground and background regions. Segmentation based EMD method is used to divide an image containing two regions of interest: object or foreground and background (Chen 2011). In the first stage, IAKFCM based image segmentation is performed and in the second stage, the EMD based refinement is done. In this second stage it uses curve evolution variation framework for segmentation. The flow fields driving the curves are based on the distribution of features in the inner and outer regions bounded by curves. Here curve evolution is based on histogram of images (Adam et al 2009). The flow fields are derived to guide the evaluation process based on EMD for measuring the dissimilarity between two histograms. It is shown in Figure 5.11.

Algorithm 5.1 EMD Refinement process

Input : Segmented image

Output: Enhanced segmented image

Itt 0:50

{ 
Segmented image pixel==1
Assign foreground

Estimate : Foreground histogram,
Increment foreground count,
Foreground cumulative histogram

Else
Estimate : Background histogram,
Increment Background count,
Background cumulative histogram
Block size=7
For
{  
Block histogram of an image
Cumulative block histogram
Block count=Cumulative block histogram/ (Block size *Block size)
}
Foreground dist = abs(Block count-Foreground count)
Background dist = abs(Block count-Background count)
If(Foreground dist< Background dist)
Foreground image
Else
Background image
End

Edge detection
Canny operator
Shape descriptor
Histogram

Figure 5.11 Algorithm for EMD Refinement process
a. Noiseless input images

b. K-means clustering

c. Fuzzy C means clustering

Figure 5.12 Continued
d. IAKFCM-EMD method

e. Human labeled segmentation

Figure 5.12 Comparative results of different segmentation method of ‘Anteater’ and ‘Butterfly’ images a) Noiseless input images b) K-means clustering c) Fuzzy C means clustering d) IAKFCM-EMD e) Human labeled segmentation
a. Noiseless input images

b. K-means clustering

c. Fuzzy C means clustering

Figure 5.13 Continued
Figure 5.13 Comparative results of different segmentation method of ‘Horse’ and ‘Elephant’ images a) Noiseless input images b) K-means clustering c) Fuzzy C means clustering d) IAKFCM-EMD e) Human labeled segmentation
Figure 5.14 Comparative results of different segmentation method of ‘Elephant’ and ‘Lizard’ images a) Noiseless input images b) K-means clustering c) Fuzzy C means clustering d) IAKFCM-EMD
Figure 5.15 Comparative results of different segmentation method of ‘Horse’ and ‘Tiger’ images a) Noiseless input images b) K-means clustering c) Fuzzy C means clustering d) IAKFCM-EMD
Figure 5.12-5.15 shows the comparative results of different segmentation methods with human-labeled segmentation. Image segmentation is an important task for computer vision applications. In this chapter, a new approach has been presented for color image segmentation based on color-texture features. The color features are extracted from statistical characteristics. Then texture feature is obtained from the fuzzy texture unit. Finally the color image is segmented using IAKFCM clustering and refinement is done based on EMD. Results obtained from the Benchmark database indicate that the proposed technique IAKFCM-EMD has produced better quantitative results than the other state-of-the-art segmentation methods recently proposed in the literature.

5.7 PERFORMANCE MEASURE

The proposed method has been adopted to segment the images on the Benchmark database. The proposed algorithm was applied to 1200 images and the output was compared with the human perceptual ground truth. The Root Mean Square Error (RMSE), Pearson correlation coefficient (r) and Structural SIMilarity (SSIM) are the three metrics used to compare image segmentation quality.

Root Mean Square Error (RMSE)

RMSE value is the difference between the segmented images and the ground truth image (Sreedevi et al 2010). The equation is shown in Equation (5.3).

\[
\text{RMSE} = \sqrt{\frac{\sum_{m=1}^{M} \sum_{n=1}^{N} (I_s(m,n) - I_g(m,n))^2}{(M+M)}}
\]  

(5.3)
**Pearson correlation coefficient**\((r)\)

Correlation – often measured as a correlation coefficient – indicates the strength and direction of a linear relationship between two variables (for example model output and observed values). A number of different coefficients are used for different situations. The best known is the Pearson product-moment correlation coefficient (also called Pearson correlation coefficient or the sample correlation coefficient), which is obtained by dividing the covariance of the two variables by the product of their standard deviations. The Pearson correlation coefficient metric is calculated on window size of 3x3 in an image. The measurement between two windows x and y of common size \(N\times N\) is represented in Equation (5.4).

\[
r = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2 \cdot (y_i - \bar{y})^2}} \tag{5.4}
\]

The correlation is +1 in the case of a perfect increasing linear relationship, and -1 in case of a decreasing linear relationship, and the values in between indicates the degree of linear relationship between for example model and observations. A correlation coefficient of 0 means the there is no linear relationship between the variables.

**Structural SIMilarity (SSIM)**

It is a method for measuring the similarity between two images. The SSIM index is a full reference metric or in other words, the measuring of image quality based on an initial uncompressed or distortion-free image as reference. The SSIM metric is calculated on window size of 3x3 in an image. The measurement between two windows x and y of common size \(N\times N\) is in Equation (5.5)
\[ SSIM_{x,y} = \frac{(2\mu_x\mu_y+\sigma_1)(2\sigma_{xy}+\sigma_2)}{(\mu_x^2+\mu_y^2+\sigma_1^2)(\sigma_x^2+\sigma_y^2+\sigma_2^2)} \] \hspace{1cm} (5.5)

- \( \mu_x \): the average of \( x \);
- \( \mu_y \): the average of \( y \);
- \( \sigma_x^2 \): the variance of \( x \);
- \( \sigma_y^2 \): the variance of \( y \);
- \( \sigma_{xy} \): the covariance of \( x \) and \( y \);
- \( C_1 = k_1L^2 \), \( C_2 = K_2L^2 \): two variables to stabilize the division with weak denominator;
- \( L \): the dynamic range of the pixel (\( 2^{\#\text{bits per pixel}} \))

\( k_1=0.01 \) and \( K_2=0.03 \) by default.

Whereas RMSE measure estimates only the perceived errors, SSIM and Pearson correlation coefficient consider image degradation as perceived change in structural information.

Structural DiSIMilarity is a distance metric derived from SSIM. It is shown in Equation (5.6)

\[ DSSIM = \frac{1-SSIM}{2} \] \hspace{1cm} (5.6)
Table 5.1 Performance Evaluation of Pearson correlation coefficient($r$) for IAKFCM-EMD Segmentation

<table>
<thead>
<tr>
<th>Sample images</th>
<th>Quantitative Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pearson correlation coefficient($r$)</td>
</tr>
<tr>
<td></td>
<td>$K$-means clustering</td>
</tr>
<tr>
<td>Anteater</td>
<td>-0.068</td>
</tr>
<tr>
<td>Butterfly</td>
<td>-0.159</td>
</tr>
<tr>
<td>Horse</td>
<td>-0.033</td>
</tr>
<tr>
<td>Elephant</td>
<td>-0.016</td>
</tr>
</tbody>
</table>

Table 5.2 Performance Evaluation of RMSE for IAKFCM-EMD Segmentation

<table>
<thead>
<tr>
<th>Sample images</th>
<th>Quantitative Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
</tr>
<tr>
<td></td>
<td>$K$-means clustering</td>
</tr>
<tr>
<td>Anteater</td>
<td>0.554</td>
</tr>
<tr>
<td>Butterfly</td>
<td>0.820</td>
</tr>
<tr>
<td>Horse</td>
<td>0.869</td>
</tr>
<tr>
<td>Elephant</td>
<td>0.807</td>
</tr>
</tbody>
</table>
Table 5.3 Performance Evaluation of SSIM and DSSIM for IAKFCM-EMD Segmentation

<table>
<thead>
<tr>
<th>Sample images</th>
<th>Quantitative Evaluation</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SSIM</td>
<td>Proposed Method</td>
<td>DSSIM</td>
<td>Proposed Method</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anteater</td>
<td>0.894</td>
<td>0.994</td>
<td><strong>0.999</strong></td>
<td>0.053</td>
<td>0.003</td>
<td><strong>0.005</strong></td>
</tr>
<tr>
<td>Butterfly</td>
<td>0.884</td>
<td>0.988</td>
<td><strong>0.999</strong></td>
<td>0.058</td>
<td>0.006</td>
<td><strong>0.005</strong></td>
</tr>
<tr>
<td>Horse</td>
<td>0.955</td>
<td>0.963</td>
<td><strong>0.975</strong></td>
<td>0.023</td>
<td>0.019</td>
<td><strong>0.013</strong></td>
</tr>
<tr>
<td>Elephant</td>
<td>0.888</td>
<td>0.960</td>
<td><strong>0.967</strong></td>
<td>0.056</td>
<td>0.020</td>
<td><strong>0.017</strong></td>
</tr>
</tbody>
</table>

Table 5.4 Time taken of IAKFCM-EMD Segmentation

<table>
<thead>
<tr>
<th>Sample images</th>
<th>Time taken (Sec)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K-means clustering</td>
<td>Fuzzy C Means clustering</td>
<td>Proposed Method (of IAKFCM-EMD)</td>
</tr>
<tr>
<td>Anteater</td>
<td>7.0</td>
<td>7.1</td>
<td><strong>6.7</strong></td>
</tr>
<tr>
<td>Butterfly</td>
<td>7.9</td>
<td>7.3</td>
<td><strong>6.9</strong></td>
</tr>
<tr>
<td>Horse</td>
<td>8.0</td>
<td>7.4</td>
<td><strong>7.1</strong></td>
</tr>
<tr>
<td>Elephant</td>
<td>8.2</td>
<td>7.7</td>
<td><strong>7.2</strong></td>
</tr>
</tbody>
</table>
Figure 5.16  Graphical representation of Pearson correlation coefficient for K-means, Fuzzy C Means and IAKFCM-EMD

Figure 5.17  Graphical representation of RMSE for K-means, Fuzzy C Means and IAKFCM-EMD
The performance evaluations for K-means clustering, Fuzzy C Means and proposed method (IAKFCM-EMD) are given in Table 5.1-5.3. The comparison is performed on the segmentation results of the proposed method on natural images obtained from Benchmark database with K-means clustering. The time taken for the above methods is depicted in Table 5.4. It is inferred that the time taken by the proposed method is small when compared to that of K-means clustering. The graphical representation of Pearson correlation coefficient, RMSE and SSIM, DSSIM and time for K-means, Fuzzy C Means (Celik & Lee 2013) and IAKFCM-EMD is shown in Figure 5.16-5.18.

5.8 SUMMARY

In this chapter shape description is obtained based on color-texture image segmentation. Shape description is one of the important features in CBIR. The statistical color features such as mean and mode are extracted from Lab color space and texture feature is determined by the extraction of
texture information from each pixel. The combined color and texture features are segmented using IAKFCM. EMD is used to eliminate the misclassified pixels. Finally the performance results obtained using IAKFCM-EMD technique compared with K-means and FCM and has proven better results in terms of Pearson correlation coefficient, RMSE, SSIM and time. Color, texture, shape and pattern based image retrieval is discussed in chapter 6.